

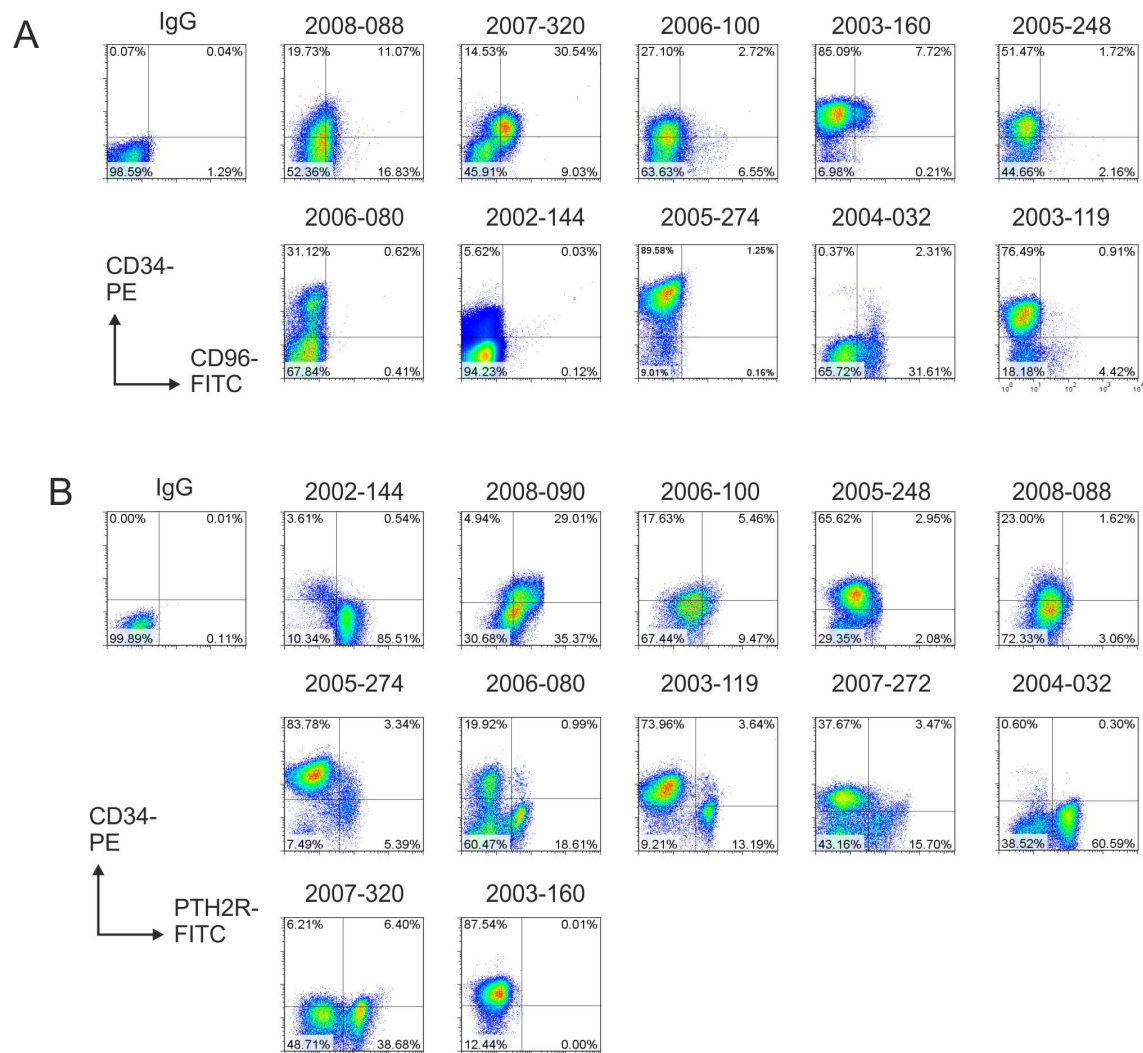
## **Supplemental Materials and Methods and Figures for the manuscript:**

### **A Proteomics and transcriptomics approach to identify leukemic stem cell markers**

Francesco Bonardi<sup>1</sup>, Fabrizia Fusetti<sup>2</sup>, Patrick Deelen<sup>1</sup>, Djoke van Gosliga<sup>1</sup>, E. Vellenga<sup>1</sup> and J.J. Schuringa<sup>1#</sup>.

<sup>1</sup>Department of Experimental Hematology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9700 RB, Groningen, The Netherlands. <sup>2</sup>Department of Biochemistry, Enzymology of membrane proteins, and Netherlands Proteomics Center, University of Groningen, The Netherlands. #to whom correspondence should be addressed, email: [j.j.schuringa@umcg.nl](mailto:j.j.schuringa@umcg.nl)

## Supplemental Figures



Supplemental figure 1. Bonardi et al 2012.

## Supplemental Materials and Methods

### Algorithm 1: Method used to select AMI CD34+ markers

**Require:**  $A \leftarrow$  expression data of AML CD34+ and BM CD34+ samples

- 1:  $A_2 \leftarrow$  genes from  $A$  that have relevant GO annotation
- 2: perform one-sided Mann-Whitney U test
- 3: select significantly upregulated genes in AML CD34+ using multivariate permutations
- 4: calculate information gain for selected genes
- 5:  $g \leftarrow$  all gene indices in  $A_2$
- 6: **while**  $g$  is not empty **do**
- 7:     find gene in  $g$  with max information gain
- 8:     add this gene to  $x$
- 9:     remove this gene from  $g$
- 10:    remove all genes from  $g$  that are correlated to selected genes in the AML CD34+ samples
- 11: **end while**